DOCUMENT RESUME

ED 052 967 SE 012 031

AUTHOR Glass, Gene V.
TITLE The Wisdom of Scientific Inquiry on Education. Science Education Information Reports, Occasional Paper Series, Science Paper 4.
INSTITUTION Colorado Univ., Boulder. Lab. of Educational Research.; ERIC Information Analysis Center for Science Education, Columbus, Ohio.
PUB DATE Jul 71
NOTE 26p.

EDRS PRICE EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS *Curriculum Development, Evaluation, Objectives, *Research Methodology, Research Needs, *Science Education

ABSTRACT

After distinguishing between elucidatory inquiry (directed toward theory and model construction for understanding and explaining phenomena) and evaluative inquiry (determining the worth of a thing), and providing nine differentiating characteristics, the author discusses the progress of educational elucidatory inquiry. He concludes that it has not "turned up any important, reliable, replicable relationships worthy of continued study...Until we get these fundamental relationships, our pretensions to elucidatory inquiry on education will be vain boasting." He feels that we should not now be attempting to build theories of teaching because ideas and techniques of teaching are changing rapidly. Under current conditions evaluative research based on the products of master teachers using basic knowledge from the social and natural sciences is more likely to contribute to the improvement of educational practice. (AL)

ED052967

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY

SCIENCE EDUCATION INFORMATION
REPORTS

OCCASIONAL PAPER SERIES - SCIENCE PAPER 4 - THE WISDOM OF SCIENTIFIC INQUIRY ON EDUCATION

by
Gene V. Glass
Laboratory of Educational Research
University of Colorado

THE OHIO STATE UNIVERSITY
ERIC Information Analysis Center
for Science and Mathematics Education
1460 West Lane Avenue
Columbus, Ohio 43221

July, 1971

SE 012 031

SCIENCE AND MATHEMATICS EDUCATION INFORMATION REPORTS

The Science and Mathematics Education Information Reports are being developed to disseminate information concerning documents analyzed at the ERIC Information Analysis Center for Science and Mathematics Education. The reports include four types of publications. Special Bibliographies are developed to announce availability of documents in selected interest areas. These bibliographies will list most significant documents that have been published in the interest area. Guides to Resource Literature for Science and Mathematics Teachers are bibliographies that identify references for the professional growth of teachers at all levels of science and mathematics teaching. Research Reviews are issued to analyze and synthesize research related to science and mathematics education over a period of several years. The Occasional Paper Series is designed to present research reviews and discussions related to specific educational topics.

The Science and Mathematics Education Information Reports will be announced in the SMAC Newsletters as they become available.

Occasional Paper Series - Science

The Occasional Paper Series (Science) is designed to review literature related to specific topics or educational programs related to the teaching and learning of science. We hope these papers will provide ideas for implementing research, suggestions for areas that are in need of research, and suggestions for research design.

Patricia E. Blosser
and
Robert W. Howe
Editors

Sponsored by the Educational Resources Information Center of the United States Office of Education and The Ohio State University.

This publication was prepared pursuant to a contract with the Office of Education, United States Department of Health, Education, and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their judgment in professional and technical matters. Points of view or opinions do not, therefore, necessarily represent official Office of Education position or policy.

TABLE OF CONTENTS

Introduction..... 1

Elucidatory and Evaluative Inquiry..... 2

The Distinction Between Elucidatory and Evaluative Inquiry..... 2

Characteristics of Inquiry..... 3

1. Motivation of the Inquirer..... 3
2. The Objective of the Search..... 3
3. Laws vs. Description..... 4
4. The Role of Explanation..... 4
5. Autonomy of the Inquiry..... 5
6. Properties of the Pheonomena Which Are Assessed..... 5
7. "Universality" of the Phenomena Studied..... 6
8. Salience of the Value Question..... 8
9. Investigative Techniques..... 8

Relationship Between Elucidatory and Evaluative Inquiry..... 9

Pretentions to Scientific Education..... 10

The Problem of Non-generalizable Knowledge..... 12

Is the Game Worth the Candle?..... 16

References

INTRODUCTION

A few years ago, Ralph Tyler assessed the state of research on science teaching and found it in need of a more theoretical, scientific perspective. My own assessment of the field now and then is diametrically opposed to Tyler's and, incidentally, to the position taken by Pella (1966) at about the time Tyler wrote on this subject. For fear that a paraphrasing of Tyler's position might distort his meaning, I shall quote him at some length. Tyler's prescription for improved research in science education was as follows:

> "Theory is the all embracing end of basic research in seeking to provide a comprehensive map of the terrain of science education. Concepts are the smaller areas which comprise the total map, or to put the metaphor in another way, the complex of science education can be understood more readily by considering the concepts as major parts of the whole, and studying these parts in greater detail than is possible with the total....The concepts and the dynamic models furnish the map which we seek in order to understand the factors involved in, and the process of science education. They form the major part of the theory....
>
> My criticism of current research is its failure to be guided by, or to produce an adequate map of the factors and processes in science education.... The outlining, elaboration, and testing of such a map seems to me to be the necessary focus of our attention if we are to improve research in this field." (Tyler, 1967-1968, p. 43.)

Tyler went on to name eleven prominent areas of "the map": 1) the objectives of science education; 2) the teaching-learning process; 3) the organization of learning experiences; 4) the outcomes of science education; 5) the student's development; 6) the development of teachers; 7) the objectives of education for science teachers; 8) the teaching-learning process of teacher education; 9) the outcomes of teacher education; 10) the organization of the teacher's learning experiences; 11) the process of change in science education.

In my opinion, we should not strive to make research on science education or education generally more scientific. Indeed, we who call ourselves educational researchers should turn away from elucidatory inquiry in all areas of education. This type of inquiry, directed toward the construction of theories or models for the understanding and explanation of phenomena, should be left to the social and natural sciences

---

Paper presented at the 1971 Convention of the National Association for Research on Science Teaching; 24 March 1971; Silver Spring, Maryland.

because it is currently unproductive in education and is a profligate expenditure of precious resources of time, money, and talent. We should turn instead toward evaluative inquiry of educational developments which are the creations of masterful teachers inspired by what reliable knowledge exists in psychology, sociology, and the other sciences concerning the educating process.

The argument depends heavily on a distinction between two types of inquiry -- elucidatory and evaluative.* Thus, it is necessary to turn attention to the definitional problem and the related problem of distinguishing between elucidatory and evaluative inquiry.

ELUCIDATORY AND EVALUATIVE INQUIRY

Elucidatory inquiry is the process of obtaining generalizable knowledge by contriving and testing claims about relationships among variables or generalizable phenomena. (Apologies to mathematics, history, philosophy, etc. for the obvious empirical social-science bias in the definition and the thinking to follow.) This knowledge results in functional or statistical relationships, models, and ultimately theories. When the results of elucidatory inquiry are combined with knowledge of particular circumstance, one obtains explanations (in the sense of Braithwaite, 1953).

Evaluative inquiry is the determination of the worth of a thing. In education it involves obtaining information to judge the worth of a program, product, or procedure. According to Scriven (1967, p. 40),

> The activity consists simply in gathering and combining of performance data with a weighted set of goal scales to yield either comparative or numerical ratings; and in the justification of (a) the data-gathering instruments, (b) the weightings, and (c) the selection of goals.

THE DISTINCTION BETWEEN ELUCIDATORY AND EVALUATIVE INQUIRY

Elucidatory and evaluative inquiry have many defining characteristics. Each is only imperfectly correlated with the tendency of informed men to call activity A "elucidatory" and activity B "evaluative" (just as clinical psychologists use "anxiety" as a construct to differentiate instances of

*The contrast between elucidation and evaluation is well-established in aesthetics. Elucidatory aesthetics is the attempt to explain what constitutes art, or more generally beauty. Evaluative aesthetics concerns the discrimination between good and bad art (what is beautiful and what is less beautiful) without explanation of why an art work is good or bad.

behavior in a way that is not perfectly reproduced by an single measure or defining characteristic). The conceptualizations of elucidatory and evaluative inquiry are enriched by the identification of any characteristic of inquiry which has a non-zero correlation with the tendency of intelligent men to speak of "elucidation" or "evaluation" when discussing a particular inquiry activity.*

CHARACTERISTICS INQUIRY

Nine characteristics of inquiry which distinguish elucidation from evaluation are recognizable.**

1. Motivation of the Inquirer

Elucidatory and evaluative inquiry appear generally to be undertaken for different reasons. The former is pursued largely to satisfy curiosity; the latter is done to contribute to the solution of a practical problem. The theory builder is intrigued; the evaluator (or at least his client) is concerned.

Although the elucidatory inquirer may believe that his work has greater long-range payoff than the evaluator's, he is no less motivated by curiosity when performing his unique function. One must be nimble to avoid becoming bogged down in the seeming paradox that the policy decision to support basic inquiry because of its ultimate practical payoff does not imply that basic inquirers are pursuing practical ends in their daily work. Scriven (1969) argues that as regards research in mathematics, practical pay-off is increased to the extent that the mathematician is convinced he is not seeking practically significant results.

2. The Objective of the Search

Elucidatory and evaluative inquiry seek different ends. The former seeks conclusions; evaluation leads to decisions (see Tukey, 1960). Cronbach and Suppes (1969, pp. 20-21) distinguished between decision-oriented and conclusion-oriented inquiry.

---

*Scriven (1958, p. 175) referred to such terms as "cluster concepts" or "correlational concepts." Such concepts, e.g., "schizophrenia," are known by their indicators all of which are imperfectly related to them.

**The next eighteen paragraphs are based on a segment of the author's paper with Blaine R. Worthen entitled "Educational Inquiry and the Practice of Education" which will appear as a chapter in Shalock, H. Del (Ed.), Frameworks for Viewing Educational Research, Development, Diffusion and Evaluation. Monmouth, Oregon; Teaching Research Division of Oregon State System of Higher Education. (In press)

> In a decision-oriented study the investigator is asked to provide information wanted by a decision-maker: a school administrator, a government policy-maker, the manager of a project to develop a new biology textbook, or the like. The decision-oriented study is a commissioned study. The decision-maker believes that he needs information to guide his actions and he poses the question to the investigator. The conclusion-oriented study, on the other hand, takes its direction from the investigator's commitments and hunches. The educational decision-maker can, at most, arouse the investigator's interest in a problem. The latter formulates his own question, usually a general one rather than a question about a particular institution. The aim is to conceptualize and understand the chosen phenomenon; a particular finding is only a means to that end. Therefore, he concentrates on persons and settings that he expects to be enlightening.

Conclusion-oriented inquiry is much like what is here referred to as elucidatory inquiry; decision-oriented inquiry typifies evaluative inquiry as well as any three words can.

3. Laws vs. Description

Closely related to the distinction between conclusion-oriented and decision-oriented are the familiar concepts of nomothetic (law-giving) and idiographic (descriptive of the particular). Elucidatory inquiry is the search for laws, i.e., statements of relationship among two or more variables or phenomena. Evaluation merely seeks to describe a particular thing with respect to one or more scales of value.

4. The Role of Explanation

Scientific explanations require scientific laws, and the disciplines related to education appear to be far from discovery of the general laws on which explanations of incidents of schooling can be based. Explanations are not the goal of evaluation. A fully proper and useful evaluation can be conducted without producing an explanation of why the product or program being evaluated is good or bad or how it operates to produce its effects.

Elucidatory inquiry is characterized by a succession of studies in which greater control ("control" in the sense of the ability to manipulate specific components of independent variables) is exercised at each stage so that relationships among variables can be determined at more fundamental levels. Science seems to be an endless search for subsurface explanations, i.e., accountings of surface phenomena in terms of relationships among variables at a more subtle, covert level. Subsurface explanations are not sought after for themselves, or else the ultimate goal of science is some personal abberation, such as the science of Lawsonomy. They are sought because greater precision, greater trust, the answer to the next "why?" always seems to be one stratum below the phenomena we see now. When ethologists find that the swallows return to Capastrono each March 19 because they are following the insects they feed on, they immediately ask

why the insects return on March 19; and so it goes. The history of learning psychology is an excellent illustration of the continual search for subsurface explanations, though countless examples could be found in any science. Psychologists have sought increasingly more fundamental explanations of learning along a path that has led through the Law of Effect, unravelings of the nature of reward, drive states, secondary reinforcement, and which leads inexorably toward brain physiology and the void beyond. Elucidatory inquiry chases subsurface explanations. However, it is usually enough for the evaluator to know that something attendant upon the installation of Harvard Project Physics (and not an extraneous, "uncontrolled" influence unrelated to the curriculum) is responsible for the valued outcomes; to give a more definite answer about what that something is would carry evaluation into analytical research.

5. Autonomy of the Inquiry

The independence and autonomy of science is so important that Kaplan (1964, pp. 3-6) wrote first of it in his classic The Conduct of Inquiry:

> It is one of the themes of this book that the various sciences, taken together, are not colonies subject to the governance of logic, methodology, philosophy of science, or any other discipline whatever, but are, and of right ought to be, free and independent. Following John Dewey, I shall refer to this declaration of scientific independence as the principle of autonomy of inquiry. It is the principle that the pursuit of truth is accountable to nothing and to no one not a part of that pursuit itself.

Not surprisingly, autonomy of inquiry proves to be an important characteristic for distinguishing elucidatory and evaluative inquiry. As was seen incidentally in the above quote from Cronbach and Suppes, evaluation is undertaken at the behest of a client who expects that particular questions will be answered. Elucidatory inquiry must be free to follow leads which pique the curiosity of those who know most intimately the aspirations of the discipline.

6. Properties of the Pheonomena Which are Assessed

Evaluation is an attempt to assess the worth of a thing, and elucidatory inquiry is an attempt to assess scientific truth. Except that truth is highly valued and worthwhile, this distinction serves fairly well to discriminate elucidatory and evaluative inquiry. The distinction can be given added meaning if "worth" is taken as synonymous with "social utility" (which is presumed to increase with improved health, happiness, life expectancy, increases in certain kinds of knowledge, etc., and decreases with increases in privation, sickness, ignorance, etc.) and if "scientific truth" is identified with two of its possible forms: 1) empirical verifiability of statements about general pheonomena with accepted methods of inquiry; 2) logical consistency of such statements. Elucidatory inquires may yield evidence of social utility, but only indirectly -- because empirical verifiability of general phenomena and logical consistency may eventually be socially useful.

In this view, all inquiry is seen as directed toward the assessment of three properties of statements about phenomena: 1) their empirical verifiability by accepted methods; 2) their logical consistency with other accepted or known facts; 3) their social utility. Most disciplined inquiry aims to assess each property in varying degrees. The definition of "theory" in Webster's Third New International Dictionary (definition 3.a (2)) is tripartite: "The coherent set of hypothetical, conceptual and pragmatic principles forming the general form of reference for a field of inquiry (as for deducing principles, formulating hypotheses for testing, undertaking actions)." The three inquiry activities in Webster's definition correspond closely to the three inquiry properties proposed here.

7. "Universality" of the Phenomena Studied

Perhaps the highest correlate of the elucidatory-evaluative distinction is the "universality"* of the phenomena studied. Elucidatory inquirers work with constructs having a currency and scope of application which make the objects one evaluates seem parochial by comparison. A psychologist experiments with "reinforcement" or "need achievement" which he regards as neither specific to geography nor to one point in time. The effect of positive reinforcement following upon the responses he observes is assumed to be a phenomenon shared by most men in most times; moreover, the number of specific instances of human behavior which are examples of the working of positive reinforcement is great. Not so with the pheonomena which educationists evaluate. A particular textbook, an organizational plan, and a film strip have a short life expectancy and may not be widely shared. However, whenever their cost or potential pay-off rises above a negligible level, they are of interest to the evaluator.

Three aspects of the "universality" of a phenomenon can be identified: 1) generality across time (Will the phenomenon -- a textbook, "Self-concept," etc. -- be of interest fifty years hence?); 2) generality across geography (Is the phenomenon of any interest to people in the next town, the next state, across the ocean?); 3) applicability to a number of specific instances of the general phenomenon (Are there many specific examples of the phenomenon being studied or is this the "one and only"?). These three features of the object of an educational inquiry can be used to classify different inquiry types as in Figure 1, page 7.

Three types of inquiry are represented in Figure 1: 1) program evaluation -- the evaluation of a complex of people, materials, and organization which make up a particular educational program; 2) product evaluation -- the evaluation of an instrument of schooling such as a book, a film, or a recorded tape; 3) elucidatory inquiry.

Program evaluation is depicted as concerned with a phenomenon (an "educational program") which has limited generalizability across time and geography; the innovative "ecology curriculum" (including instructional materials, staff, students, and other courses in the school) in the Pilger Public Schools will probably not survive the decade, is of little interest to the schools in Norfolk which have a different set of environmental

*The term is a bit grand, but it conveys a meaning that might have been distorted by many more modest labels that suggest themselves.

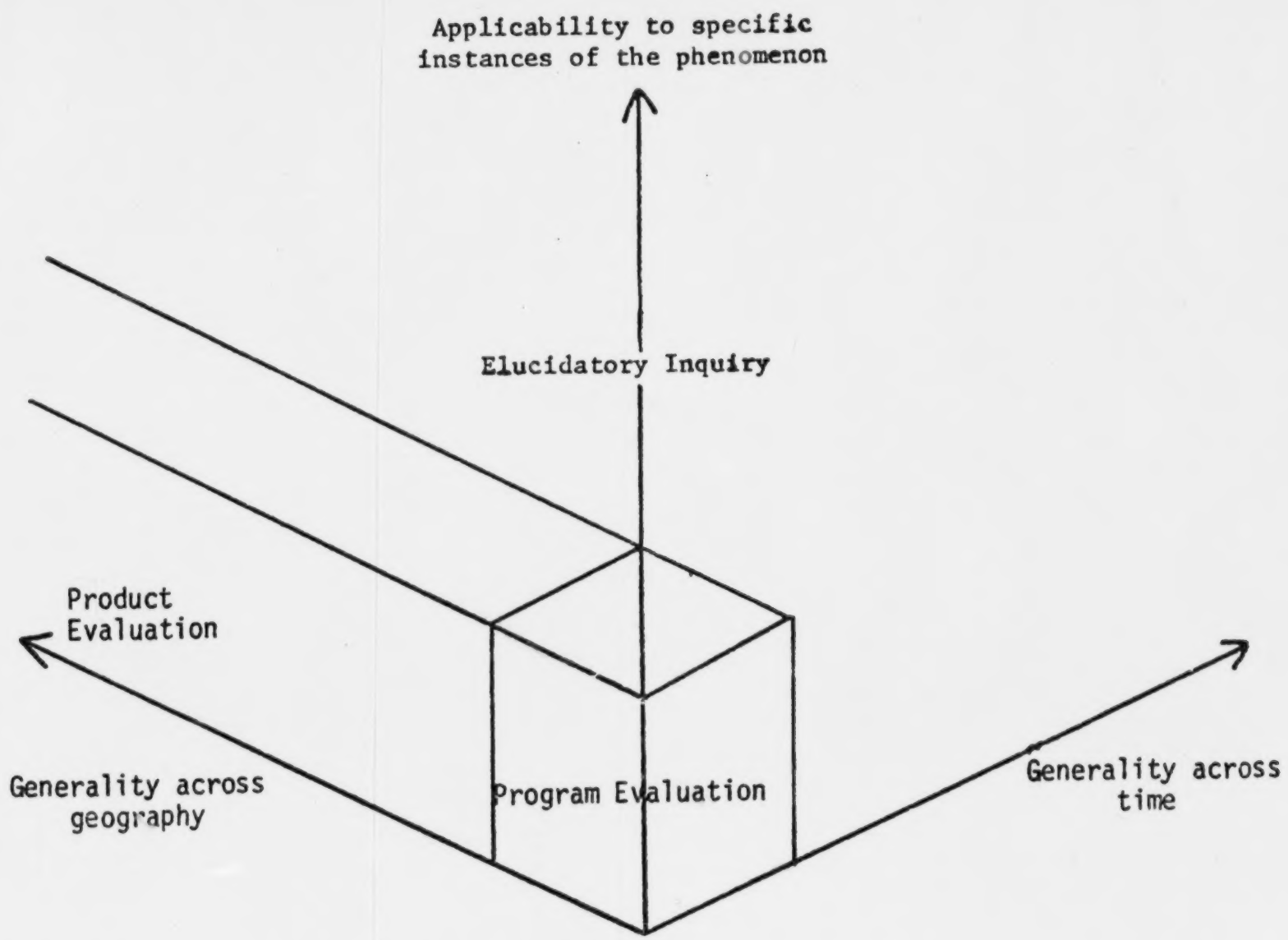

Figure 1. Three Inquiry Types Classified by the "Universality" of the Phenomenon Investigated
(The property represented by each axis is absent where the three axes meet and increases as one moves out along each axis.)

problems and instructional resources, and has little relationship to other curricula with other objectives. Product evaluation is concerned with assessing the worth of something like a new ecology textbook or an overhead projector which can be widely disseminated geographically but which similarly may not be of interest ten years hence. Furthermore, the evaluation of a particular product produces no reliable knowledge about the value of other similar -- but distinct -- products. The concepts upon which elucidatory educational inquiry is carried out are supposed to be relatively permanent, applicable to schooling nearly everywhere, and they should subsume a large number of instances of teaching and learning.

8. Salience of the Value Question

At least in theory, a value can be placed on the outcome of any inquiry, and all inquiry is directed toward the discovery of something worthwhile and useful. In evaluation, it is usually quite clear that some question of value is being addressed. Indeed, value questions are the sine qua non of evaluative inquiry and usually determine what information is sought. This is not to say that value questions are not germane in elucidatory inquiry; they are just less obvious. The acquisition of knowledge of auto-mechanics of the inculcation of "good citizenship" are clearly value-laden endeavors. The value questions in the derivation of a new oblique transformation technique in factor analysis or the investigation of the transfer of information from short-term to long-term memory are not so obvious, but they are there nonetheless. With respect to assessing the value of things, the difference between elucidatory and evaluative inquiry is one of degree, not of kind.

9. Investigative Techniques

A substantial amount of opinion has been expressed recently to the effect that elucidatory and evaluative inquiry should employ different techniques for gathering and processing data, that the methods appropriate to elucidatory inquiry -- such as comparative experimental design -- are not appropriate to evaluation, or that with respect to techniques of empirical inquiry, evaluation is a thing apart. In fact, however, there are many more similarities than differences between elucidatory and evaluative inquiry with regard to the techniques by which empirical evidence is collated and judged to be sound. As Stake and Denny (1969, p. 374) indicated: "The distinction between research and evaluation can be overstated as well as understated....Researchers and evaluators work within the same inquiry paradigm....(training programs for) both must include skill development in general educational research methodology."

Hemphill (1969, p. 220) expressed the same opinion when he wrote: "The consequence of the differences between the proper function of evaluation studies and research studies is not to be found in differences in the subject of interest or in the methods of inquiry of the researcher and of the evaluator."

RELATIONSHIP BETWEEN ELUCIDATORY AND EVALUATIVE INQUIRY

Evaluation borrows inquiry techniques and knowledge for recognizing value from basic sciences and contributes little in return. Methodological research in the social sciences produces the technologies of data collection and analysis that are so important to empirical educational evaluation. In addition, knowledge produced by basic elucidatory inquiry is often critical in determining whether a particular finding from an evaluative study counts as good or bad. For example, the evaluative meaning given to the finding that a health curriculum decreased the incidence of cigarette smoking among teenagers is dependent upon the extensive medical research which established a casual link between cigarette smoking and cancer and coronary disease.

The view of the relationship between elucidatory and evaluative inquiry presented here is parasitic; some see the two living symbiotically:

> To some extent evaluative research may offer a bridge between 'pure' and 'applied' research. Evaluation may be viewed as a field test of the validity of cause-effect hypotheses in basic science whether these be in the field of biology (i.e., medicine) or sociology (i.e., social work). Action programs in any professional field should be based upon the best available scientific knowledge and theory of that field. As such, evaluations of the success or failure of these programs are intimately tied to the proof or disproof of such knowledge. Since such a knowledge base is the foundation of any action program, the evaluation research worker who approaches his task in the spirit of testing some theoretical proposition rather than a set of administrative practices will in the long run make the most significant contribution to the program development. (Suchman, 1967, p. 1970)

Suchman saw "action programs" as based on scientific knowledge and theory; I see a more tenuous link between (a) what can be conceived of abstractly and established empirically in the laboratory and (b) what can be implemented in the field. There can hardly be said to exist an "intimate" connection between the success of a program in the field and some theory or hypotheses from a basic discipline. The tribal medicine man may be effective for all of the wrong reasons, and a prototype turbine automobile may fail even though it is consistent in all particulars with physical theory.

One need not be able to explain phenomena to evaluate them. We can know how well or how poorly without knowing why. Understanding or elucidation is required neither in summative nor formative evaluation. We can perfectly well make summative judgments -- and we frequently do -- without understanding why one of the alternatives being evaluated scored highest on some weighted scale of value. In formative evaluation, observations of the quality of performance are made for the purpose of

improving the thing that we are developing. It may seem that one can't very well make an improvement in the system unless he knows why performance was unacceptable. Otherwise, his responses to data on substandard performance would be random stabs in the dark with little chance of success. If tests scores indicate the prevelance of ignorance among tenth graders after studying a program unit on the relationship between birth rates, death rates, and population growth, the curriculum developers will act quite differently if they suspect the failure is due to the lack of practice in working problems rather than the complexity of the mathematics used to explain the relationship. Surely then, formative evaluation requires some greater knowledge of why performance is acceptable or unacceptable than does summative evaluation. But this greater knowledge is knowledge of particulars not codified general knowledge in the form of an abstract system of laws. We know why many things are as they are, even why some things are good and others bad; and all of this knowledge is of particulars, it is of no general significance. My car continually loses front-end alignment because of the accident that sprung the frame; my electric razor works poorly because I dropped it on the bathroom floor. Both statements are explanations, but that doesn't make them of general interest. Thus we pursue both formative and summative evaluation without the need to explain outcomes by reference to a system of general laws or relationships among variables.

PRETENTIONS TO SCIENTIFIC EDUCATION

The conclusion forces itself upon one that elucidatory inquiry on education over that past eighty years (from Joseph M. Rice to the greenest Ph.D. just now taking his final orals) has been a failure for the most part. The areas that I would regard as distinctly educational research and not highly dependent upon research in one of the behavioral or social sciences are such as the following: teacher behavior, teacher competence, classroom interaction, guidance theory, counseling, tests and measurement, media of instruction (e.g., television and movies), the organization of teaching (e.g., team teaching, cross-age teaching, and differentiated staffing), class size research, and many of the prominent areas of Ralph Tyler's (1967-68, p. 44) map of science education (the outcomes of science education, student development -- perhaps as opposed to human development in general -- the development of teachers, the objective of education for teachers, the teaching-learning process of teacher education, and the process of change in education, etc.).

I am not under the mistaken notion that scientific inquiry must result in errorless functional relationships among variables. Probablistic laws are the order of the day throughout the sciences (physical, natural, or social). Our codified bodies of knowledge in any area are actually tendency statements about the occurrence of phenomenon A tending to be associated with the occurrence of phenomenon B. But educational research has hardly produced even tendency statements about tendency statements, i.e., that one has observed a tendency for A to tend to be associated with B sometimes and under some conditions.

Allan Mazur (1968) gave a wry assessment of the status of sociology vis a vis the more established physical and natural sciences. Mazur used as a touchstone for determining when a body of inquiry had become a science whether the discipline strikes the vast majority of people-in-general as profound. A science, he maintained, "is a body of theoretical knowledge that is not trivial. The scientist must have better theories than the layman or he's really not a scientist at all. If physics is a science because it is empirical and theoretical, it is also a science because physicists' theories about the physical world work better than the non-physicists' theories." "An empirical theoretically connected body of knowledge is science only when the people who know about the theories know more about the real world than the people who do not know the theories." At that point, Mazur asked whether sociology is a science, and answered, No. Perhaps the same frame of mind led a well-known philosopher of science to remark at lunch one day that John McDonald, author of the Travis McGee detective story series, is a better sociologist than most contemporary sociologists. Walter Lippman, who seldom revealed his academic credentials (which as a student of William James and George Santyana were considerable), made the same point several years before Mazur about the psychology of the 1920's:

> "We can be confident that on the whole a good meteorologist can tell us more about the weather than even the most weather-wise old sea captain. But we cannot have that kind of confidence in even the best of psychologists. Indeed, an acquaintance with psychologists will, I think, compel anyone to admit that, if they are good psychologists, they are almost certain to possess a gift of insight which is unaccounted for by their technical apparatus. Doubtless it is true that in all the sciences the difference between a good scientist and a poor one comes down at last, after all the technical and theoretical procedure has been learned, to some sort of residual flair for the realities of the subject. But in the study of human nature that residual flair, which seems to be composed of intuition, commonsense, and unconsciously deposited experience, plays a much greater role than it does in the more advanced sciences." (Lippman, 1929, p. 162)

The "science of education" certainly fails in its pretentions to scientific status if we use Mazur's criterion. There is probably far more knowledge (about how to manage and promote learning in the classroom) in the nervous systems of ten excellent teachers than an average teacher can distill from all the educational research journals in existence. Hawkins (1966) wrote that in teaching-learning "...the best practice excels the best theory in quite essential ways...."

Teacher behavior studies provide a good illustration of the unproductive nature of most elucidatory educational inquiry. The analogy is uncomplimentary, however true it is, that most teacher behavior

studies, indeed, even the best of them, have found relationships no stronger than the evidence for the ability of graphanalysts to assess personality through handwriting samples. The data in Table 1, Page 13, show a typical graphanalyst's ability to rate personality traits from handwriting samples to be nearly as good as many judge's ability to rate teacher behavior characteristics and the subsequent relation of these characteristics to the gains made by students on achievement tests. In Table 2, Page 14, correlations are shown between teacher behaviors (based on the works of Bellack, Smith, Taba, Flanders, etc.) and pupil achievement from one of the best studies ever performed in this area. The correlations are not indicative that instructional researchers' understanding of these phenomena is conspicuously superior to graphanalysts insight into personality.

The study of teacher behavior and its relation to student performance may be even worse off than my impertinent analogy would indicate. Barak Rosenshine has recently reviewed the experimental literature growing out of the correlational teacher behavior studies and concluded that the results of these experimental studies are greatly discouraging. When a variable once shown correlationally to have a weak or moderate relationship to student performance as manipulated experimentally, in almost all cases no influence on student performance was observed. Such is the residue of some hundreds of thousands of man-hours of elucidatory inquiry about the dominant personality in the classroom.

Certainly one of the most important factors militating against successful elucidatory inquiry about education is the enormous complexity of the system educational researchers seek to understand. This reason is often advanced by educational researchers themselves, and in advancing it they bring down upon themselves accusations that they are merely making excuses for lack of diligence and insight. I certainly do not make such an ungenerous assessment of this explanation for the failure of a great deal of educational research. The system educational researchers study is as complex as any that science has ever dealt with successfully; a lake or the human heart is trivial by comparison.

But the fact remains that elucidatory inquiry on education simply does not seem to have turned up any important, reliable, replicable relationships worthy of continued study. Several of these will be needed before any body of theory or even a modestly complicated model of education can be constructed. Until we get these fundamental relationships our pretentions to elucidatory inquiry on education will be vain boasting.

THE PROBLEM OF NON-GENERALIZABLE KNOWLEDGE

Psychology has the phenomenon of reinforcement which works fairly reliably. The fascinating thing about Piaget's work on cognition is that it demonstrates such a stable, replicable phenomenon: Take any three-year-old off the streets of Cleveland or out of the rain forests of Obidos, and he won't conserve mass. But the phenomena and facts that educational research has unearthed are so fragile as to scarcely deserve the name "fact."

Table 1

Correlations Among Ratings by a Handwriting Analyst and Two Counseling Psychologists

(n = 22)*

	Trait				
	Frankness	Self-confidence	Clarity of goals	Emotional Control	Rigidity
Counseler 1 vs. Counseler 2	.36	.57	.74	.21	.55
Handw. Analyst vs. Counseler 1	.02	.21	.40	-.46	-.36
Handw. Analyst vs. Counseler 2	-.28	.19	.72	.23	-.12
Handw. Analyst vs. Counselers combined	-.11	.22	.53	-.05	-.27

*From Kimmel, Douglas, and Wertheimer, Michael, "Personality Ratings Based on Handwriting Analysis and Clinical Judgment: A Correlational Study," Journal of Projective Techniques and Personality Assessment, 30: 177-178; 1966.

Table 2*
Correlations Between Teacher Behavior Variables and Pupil Achievement**

(n = 17 classes)

Teacher Behavior	Correlation with Pupil Achievement	Teacher Behavior	Correlation with Pupil Achievement
Teacher talk	-.09	Name child	.24
Teacher utterances	.35	Redirects questions	.54
Structuring (total)	-.13	Reciprocates to extend	.20
Teacher info after questions	-.52	Reciprocates to lift	-.20
Terminal structuring	.41	Affirmative and negative comment	-.14
Utterances with one question	.54	Indefinite and complex comment	.35
Utterances with two questions	-.42	Reflecting comment	.06
Utterances with more questions	-.43	Thanks and praise	.49
Closed questions	.31	Managerial comment	-.22
Open questions	-.08	Challenging comment	-.38
Alternative subsequent questions	-.40	Repetition of response	.17
		Monologues	-.01
		Recapitulation (2)	.18
		Recapitulation (3)	-.08
		Review at end of lesson	.67

*Adapted from Wright, C. J. and Nuthall, G., "Relationships between Teacher Behavior and Pupil Achievement in Three Experimental Elementary Science Lessons," American Educational Research Journal, 7: 477-491, 1970.

**Mean achievement for a class was measured as the residual of the mean posttest score for the class on an achievement test from the regression plane determined by mean IQ and pretest on science concepts for the class. Note: the largest r in the table, .67 with "review at end of lesson," is clouded by the possibility that amount of reviewing at the end of lesson was influenced by the teachers' perception of what the students appear to learn during the lesson, rather than vice versa. This objection can also be raised with regard to the correlation of .41 between pupil achievement and "terminal structuring"; the teachers' tendency to structure the subject at the end of the lesson might have been caused (rather than be the cause of) by how successfully pupils learned the lesson which was later structured.

Interactions predominate in the accumulated wisdom of educational research. A relationship among variables appears and disappears so often as a function of extraneous conditions that one genuinely never knows when to expect to see the relationship.

It has been observed that a young man seeking to make his reputation as a scholar could hardly do better than choose educational research. Current findings in the field are so interactive and ungeneralizable that a novice need only stake out a small area of interest and apply himself diligently to studying it. He will surely rise to world-wide eminence confident that his colleagues in related areas will never devise models or theories sufficiently general either to subsume his work or to prove it wrong.

The most useful scientific laws are those of physics, most of which are learned firsthand and nearly unconsciously early in life by everyone. These basic physical laws are nearly perfectly generalizable; there is no evidence that they have ever been suspended on the surface of the earth at any time in history. An astronaut in his space-capsule experiences the repeal of certain physical laws and must laboriously learn a new set.

By contrast, the laws of the social and behavioral sciences are of extremely limited generality. These laws are highly interactive with factors such as geography, time, culture, and characteristics of persons. If physical laws were as limited in generality as the laws so far discovered by social scientists, we would hesitantly creep out of bed each morning not knowing whether we would float to the ceiling or crash to the floor. If physical laws were as erratic as the "laws" governing the educational system, we wouldn't dare to get out of bed.

Consider one of the most interesting areas of elucidatory inquiry in education in recent years: the work of Rothkopf, Frase and Anderson and others on mathemagenic behavior, in particular the control of attention. These researchers have demonstrated the striking effects on learning of the control of attention through prompts in programed instruction and adjunct questions in textual materials. On the latter question, it has been demonstrated that questions preceding paragraphs to be read cue the learner's attention to the answers to the questions and away from incidental material. Post-paragraph questions cannot cue specific passages in a paragraph and thus result in greater acquisition of information. In programed instruction, over-prompting of stimulus and response in associative learning has been shown to cause inattention to the stimulus-response pairing resulting in poor learning.

The conclusions these researchers have reached concerning the cueing and control of attentional processes are as well understood and lawful as basic educational research can hope for. Yet further work in this area has shown the facilitative effect to be highly interactive with such factors as positioning and pacing of questions, the difficulty of the material to be learned, the propinquity of question and answer, whether questions are general or specific, whether answers to questions are common or uncommon words, whether subjects can turn back in a text for review, habituation to the adjunct questions, the motivation of the learner, amount of rewards for

learning, meaningfulness of material to be learned, immediate or delayed retention, the mode (visual or auditory) or presentation of the material, etc. The mathemagenic effect appears and disappears as a function of these mediating conditions even in the scientists' own make-believe world of ceteris paribus, far away from the welter of variables in a typical classroom.

J. M. Stephens in The Process of Schooling (1967) surveyed a huge portion of the educational research literature. He sampled the more trustworthy regions of that literature and came up with a distressing showing of irregularity in the findings of educational researchers for the past sixty years. Stephens' argument is that in a dozen or more areas of the teaching-learning domain, the studies investigating a certain relationship between variables divide about equally between positive and negative findings. I interpret Stephens's summary to show again that relationships among variables in education are highly interactive. To put the matter in slightly different language, the relationships which educational researchers investigate are highly context dependent. A relationship between televised instruction and student learning appears here because of administrative support and fails to appear there because of teacher attitude. In other words, the relationships among educational phenomena are very likely mediated by a huge number of influences, most of which we have yet to conceptualize properly, let alone measure validly and reliably.

If educational researchers would strive for generalizability of findings, by attempting to relieve context dependence, I am convinced that their inquiries would be pushed further and further toward the psychologists' laboratory. If elucidatory inquiry in education could be made less and less context dependent, it would look more and more like basic psychological inquiry.

IS THE GAME WORTH THE CANDLE?

We are heading in this paper straight toward the question of whether education should persevere in attempts to build a science of education, that is, in attempts to build codified bodies of knowledge leading to models and ultimately theories about the teaching-learning and broader educational process. In a recent interview concerned with the prospects of educational research, the philosopher of science, Thomas Kuhn, seemed to answer the question just posed in the negative: "I'm not sure that there can now be such a thing as really productive educational research. It is not clear that one yet has the conceptual research categories, research tools, and properly selected problems that will lead to increased understanding of the educational process. There is a general assumption that if you've got a big problem, the way to solve it is by the application of science. All you have to do is call on the right people and put enough money in and in a matter of a few years, you will have it. But it doesn't work that way, and it never will." (See Dershimer (1970, p. 79).)

I interpret Kuhn's remark to imply that the construction of a science of education is not impossible, but it is scarcely the simple matter of a redoubling of effort and faith in the application of the scientific method. The opinion is frequently expressed that educational research can never be basic research because education is a practice unlike psychology and biology which are "basic." Such reasoning is mere word play, not unlike arguing that basic research in physics is impossible because a carpenter practices physics when he sets a screw (inclined plane) or pulls a nail (lever and fulcrum). Ebel (1967, p. 81) gave as one of his reasons that "basic research in education can promise very little improvement in the process of education" that "the process of education is not a natural phenomenon of the kind that has sometimes rewarded scientific study in astronomy, physics, chemistry, geology, and biology." Ebel's selection of examples deftly excluded all of the social sciences. One wonders whether he would regard them as studies of natural phenomena. The English jurist Glanville Williams dispensed with arguments from nature against contraception by saying that "the statement that an act is unnatural coming from a moralist means little more than that he doesn't like it." Ebel's use of the word appears to mean little more than that he would prefer to exclude education from basic inquiry. I do not. I believe that we could pursue with success the construction of a science of education. This pursuit could culminate in non-trivial theories with explantory and predictive power. But the question before us now is should we? Should we do it now and is it worth the price?

In my opinion, we should not. Attempts to build a theory to science teaching, for example, face the problem of a rapidly changing object of study. The phenomena that physicists have attempted to explain have a certain timeless quality about them: force, velocity, mass, etc. To be sure physicists probe new phenomena decade after decade. However, they have confidence that the stuff of which physical studies are made, will be around for a long time. In the social sciences one finds far less stability in the subject matter. Inquiries can become obsolete; circumstances of society change. John Kenneth Galbraith pointed out quite clearly in The Affluent Society how industrialization and more sophisticated technology upset classical relationships among supply and demand and changed the function of the market. The change in the phenomena of interest to educational researchers is even more chaotic because educators pick up and discard one new idea after another. To the extent that researchers have extracted something stable and of interest out of this welter of educational practices, they look and act more like learning psychologists than like scientists studying the educational system. Published educational research has an aroma of Chippendale about it. It is generally dated as compared with the concerns of contemporary schools (as witness research on handwriting, "study hall," instructional radio, progressivism, teaching machines, attendance, et al., which once seemed so timely and today seem as quaint as the Packard, Allen's Alley, or Bill Haley and the Comets). While the researcher is groping for understanding of the effects of televised instruction in large groups, the schools will long since have moved on to dual audio television or video-type casettes embedded in individualized instruction.

The problems of obsolescence of educational research findings would be aggrevated if researchers followed some of the more cogent suggestions for building a science of teaching-learning. Jack Easley, a former colleague of

mine at the University of Illinois, maintained that there was some hope for a modest science of education as long as the models or theory included learner, teacher, and subject matter as elements. He felt that attempts to theorize in general without regard for the structure of the subject matter to be taught would flounder when applied to particular instances of teaching. Thus we might successfully pursue the construction of a theory of teaching the concept of natural selection to post-adolescents, but anything as general as a theory of teaching science to human beings was certain to fail. I regard Easley's position as wise counsel. Its implications for the possibility of building a science of education are dire, however. For what is more fluid and changing than the concepts and subject matter that each new generation of educators feels must be taught in the classroom. Recently a biologist in charge of one of those massive undergraduate general biology courses told me that in reviewing a 1940's vintage college biology text he discovered virtually no content he would want to teach today.

It can be argued -- and not easily refuted -- that educational change is so chaotic because innovations are not based on scientific knowledge. This position implies that the pace of change would diminish if education rested on a foundation of reliable basic knowledge of teaching-learning. A little knowledge from basic elucidatory inquiry might stabilize the educational system and thus permit more extensive elucidatory inquiry. This position probably attributes too much power to scientific knowledge as a controlling force of a social institution.

My pessimism for elucidatory inquiry in education stems not so much from a lack of faith in our intelligence or ability to pursue scientific inquiry about anything we choose, as it does from an assessment of the resources one could reasonably assume that society could spare for the task of building a science of education. In his proposal for the National Institute of Education, Roger Levien reported on the manpower available for research, development and innovation in education. In fiscal year 1968 the total effort from all sources (public, private, governmental, etc.) for research on education was 1,930 man years. The definition of research in that report included both basic and applied, and hence was construed more broadly than I have used the term "elucidatory inquiry" in this paper. We can safely conclude, then, that as recently as three years ago the total effort directed toward the construction of a systematic body of knowledge about education amounted to substantially less than four thousand persons in the United States working half days. By contrast, in the same year, 1968, nearly 15,000 full-time researchers were grappling with the problems of agricultural research, some fraction of them no doubt building a science of agronomy. In the health sciences nearly 60,000 man years were expended on research and development in fiscal year 1968. The poverty of the educational research effort is more than we might have expected. The construction of a science of education might drag on for many decades if it depended solely on the efforts of 1,930 full-time "educational scientists." In short, the resources for inquiry of any type, elucidatory or evaluative, in education are severely limited, will probably remain so for some time, and perhaps ought not be squandered on such a dubious endeavor as building a science of education. Scriven (1969, p. 19) remarked in the context of a discussion of pure and applied research in psychology that "it is much better to move into theory exactly where and only when obliged to by the combination of data and

needs that define our tasks. Speculation in the absence of a clarification of these parameters is too often merely idle, the kind of irresponsible gambling with society's resources that is lauded in cheap histories of science and was once a rich amateur's perogative."

We can afford not to seek a theory of education at this time simply because men have always lived without complete understanding. Indeed it is usually better and safer at the level of the whole society to know how well than to know why, if we must choose one or the other. Men do many things better than they ought to, e.g., football, trumpet playing. We can do many things better than generalizable codified knowledge about those things would seem to permit. A great deal of our knowledge is inexpressible, yet it exists and is important nonetheless. Anthropologists have discovered primitive societies that are ignorant of the relationship between sexual intercourse and pregnancy, yet presumably the existence of these societies is no more precarious for that ignorance.

Persons with a taste for empirical, disciplined inquiry on science teaching have more serious business to undertake than elucidatory inquiry on education. Their attentions ought to be focused on educational development and evaluative inquiry. Such developments in science education and elsewhere might best be regarded as bearing a heuristic relationship to elucidatory inquiry in the social sciences, particularly the science of human learning. I have developed this notion of a heuristic relationship between basic and applied knowledge to some length elsewhere and space does not permit repeating it here (see Glass, 1971). Perhaps it is sufficient to say only that the word "heuristic" is used in the literal sense, namely suggesting or stimulating empirical research. Basic knowledge will never prescribe a particular practice, but it can stimulate creative minds who have some understanding of that basic knowledge to develop materials and practices based loosely on it. Basic knowledge out of psychology and other social and natural sciences ought to be mediated through the minds of the one in ten thousand masterful teachers (the Gattegnoes, Papiis, and Ashton-Warners) currently at large who would create curriculum materials under the inspiration of this knowledge. My model of educational development involves inspiring these creative geniuses with what reliable knowledge exists in the basic sciences that is conceivably relevant to education, allowing these teachers to create developments, and then subjecting their creations to the reality testing of evaluation. Unfortunately there is not space here to develop these ideas fully. Some of the flavor of my position can be sensed by comparing and contrasting it with positions that have been published. My assessment of the pay-off elucidatory inquiry in education agrees totally with Atkin's (1967-68) assessment of "behavioral science" research on science teaching. I am more taken with the "engineering model" of development than he, and I resonate to the tones of his eloquent description of the "naturalistic model" of research and development, though the resonance is dampened somewhat by the same intolerance of ambiguity that makes the "engineering model" attractive to me.

There is much work to be done in the development of the evaluation methodology in education. An enormous amount of work remains to be done in the measurement of human behavior. I am encouraged by the amount of effort in that direction that researchers on science teaching have made.

The results, as they appear in the Journal of Research in Science Teaching, are increasingly impressive. For the next several decades then, I am placing my money on evaluative inquiry (with a heavy emphasis on measurement) which is integrated with the process of educational development in which master teachers play an important role. As for elucidatory inquiry about science teaching or education generally, I am far less optimistic.

REFERENCES

Atkin, J. Myon. "Research Styles in Science Education." Journal of Research in Science Teaching, 5: 338-345, 1967-68.

Braithwaite, Richard B. Scientific Explanation. Cambridge, England: Cambridge University Press, 1953.

Cronbach, Lee J. and Suppes, Patrick. Research for Tomorrow's Schools: Disciplines Inquiry for Education. New York: MacMillan, 1969.

Dershimer, Richard A. "The Educational Research Community: Its Communication and Social Structure." Bureau of Research, U. S. Office of Education, Project No. 8 - 0751. April, 1970.

Ebel, Robert L. "Some Limitations of Basic Research in Education." Phi Delta Kappan, pp. 81-84, October, 1967.

Glass, Gene V. "Educational Knowledge Use." Educational Forum, 1971 (in press).

Hawkins, David. "Learning the Unteachable." In Shulman, L. and Keislar, E. (eds.), Learning by Discovery: A Critical Appraisal. Chicago: Rand McNally, 1966.

Hemphill, John K. "The Relationship between Research and Evaluation Studies." In Ralph W. Tyler (ed.), Educational Evaluation: New Roles, New Means. The 68th Yearbook of the National Society for the Study of Education, Part II. Chicago, National Society for the Study of Education, 1969.

Kaplan, Abraham. The Conduct of Inquiry. San Francisco: Chandler, 1964.

Lippman, Walter. A Preface to Morals. New York: MacMillan, 1929.

Mazur, Allan. "The Littlest Science." The American Sociologist, 3: 195-200, 1968.

Pella, M. O. "A Structure for Science Education." Journal of Research in Science Teaching, 4: 250-252, 1966.

Suchman, Edward A. Evaluative Research. New York: Russell Sage Foundation, 1967.

Scriven, Michael. "Definitions, Explanations, and Theories." In Herbert Feigel, Michael Scriven, and G. Maxwell (eds.), Minnesota Studies in the Philosophy of Science, Vol II. Minneapolis, Minnesota: University of Minnesota Press, 1958.

Scriven, Michael. "Psychology Without a Paradigm." In Louis Berger (ed.), Clinical-cognitive Psychology: Models and Integrations. Englewood Cliffs, New Jersey: Prentice-Hall, Incorporated, 1969.

Scriven, Michael. "The Methodology of Evaluation." In R. E. Stake (ed.), AERA Monograph Series on Curriculum Evaluation, No. 1. Chicago: Rand McNally, 1967.

Stake, Robert E. and Denny, Terry. "Needed Concepts and Techniques for Utilizing More Fully the Potential of Evaluation." In Ralph W. Tyler (ed.), Educational Evaluation: New Roles, New Means. The 68th Yearbook of the National Society for the Study of Education, Part II. Chicago: National Society for the Study of Education 1969.

Stephens, J. M. The Process of Schooling: A Psychological Examination. New York: Holt, Rinehart & Winston, 1967.

Tukey, John W. "Conclusions vs. Decisions." Technometrics, 2: 423-433, 1960.

Tyler, Ralph W. "Resources, Models and Theory in the Improvement of Research in Science Education." Journal of Research in Science Teaching, 5: 43-51, 1967-68.